Commentary

New Toxicology Tools and the Emerging Paradigm Shift in Environmental Health Decision-Making

Gary L. Ginsberg,[1]† Kristi Pullen Fedinick,[2] Gina M. Solomon,[3] Kevin C. Elliott,[4,5,6] John J. Vandenberg,[7] Stan Barone Jr.,[8] and John R. Bucher[9]

[1]Yale School of Public Health, Yale University, New Haven, CT
[2]Natural Resources Defense Council, Washington, DC
[3]University of California, San Francisco School of Medicine, San Francisco, California
[4]Department of Fisheries and Wildlife, Michigan State University, East Lansing, Michigan
[5]Lyman Briggs College, Michigan State University, East Lansing, Michigan
[6]Department of Philosophy, Michigan State University, East Lansing, Michigan
[7]Center for Public Health and Environmental Assessment, U.S. Environmental Protection Agency (EPA), Research Triangle Park, North Carolina
[8]Office of Chemical Safety and Pollution Prevention, U.S. EPA, Washington, DC
[9]National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina

BACKGROUND: Numerous types of rapid toxicity or exposure assays and platforms are providing information relevant to human hazard and exposure identification. They offer the promise of aiding decision-making in a variety of contexts including the regulatory management of chemicals, evaluation of products and environmental media, and emergency response. There is a need to consider both the scientific validity of the new methods and the values applied to a given decision using this new information to ensure that the new methods are employed in ways that enhance public health and environmental protection. In 2018, a National Academies of Sciences, Engineering, and Medicine (NASEM) workshop examined both the toxicological and societal aspects of this challenge.

OBJECTIVES: Our objectives were to explore the challenges of adopting new data streams into regulatory decision-making and highlight the need to align new methods with the information and confidence needs of the decision contexts in which the data may be applied.

METHODS: We go beyond the NASEM workshop to further explore the requirements of different decision contexts. We also call for the new methods to be applied in a manner consistent with the core values of public health and environmental protection. We use the case examples presented in the NASEM workshop to illustrate a range of decision contexts that have applied or could benefit from these new data streams. Organizers of the NASEM workshop came together to further evaluate the main themes from the workshop and develop a joint assessment of the critical needs for improved use of emerging toxicology tools in decision-making. We have drawn from our own experience and individual decision or research contexts as well as from the case studies and panel discussions from the workshop to inform our assessment.

DISCUSSION: Many of the statutes that regulate chemicals in the environment place a high priority on the protection of public health and the environment. Moving away from the sole reliance on traditional approaches and information sources used in hazard, exposure, and risk assessment, toward the more expansive use of rapidly acquired chemical information via *in vitro*, *in silico*, and targeted testing strategies will require careful consideration of the information needed and values considerations associated with a particular decision. In this commentary, we explore the ability and feasibility of using emerging data streams, particularly those that allow for the rapid testing of a large number of chemicals across numerous biological targets, to shift the chemical testing paradigm to one in which potentially harmful chemicals are more rapidly identified, prioritized, and addressed. Such a paradigm shift could ultimately save financial and natural resources while ensuring and preserving the protection of public health. https://doi.org/10.1289/EHP4745

Introduction

From high-throughput cell-based *in vitro* studies, to tissues-on-a-chip, to genetically diverse rodent strains and environment-wide association studies, newer approaches have led to new understandings about the potential effects of chemical exposures on human health and the environment. These new approaches are faster and less expensive than traditional animal toxicity testing approaches and are becoming increasingly relevant to public health and environmental decisions. However, many questions

remain about whether and how to adopt new methods and about the values and value judgments that should inform a paradigm shift toward newer approaches. An overriding question is how to build confidence in the development and use of new data streams so that they can become the basis for the wide array of research, policy, and regulatory decisions facing the environmental health field.

This challenge was the central theme of a 20–22 November 2017 NASEM workshop titled "Understanding Pathways to a Paradigm Shift in Toxicity Testing and Decision Making" (NASEM 2018). The purpose of this meeting was to explore how scientists, policy makers, risk assessors, and regulators incorporate new science into their decisions and to raise awareness about the questions that need to be addressed in order to build confidence in the data emerging from new approaches and their use in public health protection. The conference also included lessons from the social sciences about decision-making processes and what is required to build confidence during times of change. The workshop was not focused on consensus building, but it provided a well-rounded discussion of the issues involved in using emerging data streams, including perspectives from researchers, regulators, and a variety of regulated and public interest stakeholders. A summary of the workshop is available (NASEM 2018).

In this commentary, we further explore the issues identified in the NASEM Paradigm Shift workshop and highlight what we

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

believe to be key ingredients needed for the expanded use of emerging techniques in decision-making. We utilize case studies from the workshop to illustrate how a range of decision contexts can benefit from the new data streams and provide an additional perspective on the need for these methods to be applied in a manner that advances the goals of public health and environmental protection.

Background

Whole animal toxicity testing has historically been the gold standard of safety or risk evaluation across a range of regulatory domains, including those for pharmaceuticals, industrial chemicals, pesticides, and environmental contaminants. For a few environmental chemicals, such as ozone, controlled human exposure studies and human epidemiologic studies have been most informative (U.S. EPA 2015b). Although these classical toxicology and epidemiology approaches have been widely used to identify chemical hazards and to provide dose–response data for risk assessments, the progress in evaluating the enormous universe of chemicals to which humans, animals, and ecosystems are exposed has been limited. In this commentary, we support the contention that the inclusion of methodologically diverse data streams could rapidly increase our knowledge base and predictive skills, improve decision-making, and enhance the ability of regulatory agencies to protect human health and the environment.

Methods

As organizers of the NASEM Paradigm Shift workshop, we saw a need for a broad and robust discussion of the issues surrounding the use of emerging toxicology information in decision-making. We evaluated the themes and technical information presented at the workshop to further define key challenges and potential solutions for making emerging data streams more useful and of

higher confidence for decision-makers in a variety of public health and environmental contexts. We identified and developed three workshop themes: *a*) the importance of decision context, *b*) approaches for instilling confidence in the new methods, and *c*) the need to consider value judgments about the level of precaution and public health protection associated with the use of these new data in decision-making. Although the workshop was not a consensus meeting, this commentary represents a unified perspective from the authors and is not meant to represent the views of the workshop participants, full organizing committee, or NASEM.

Discussion

In the following sections, we explore several major themes from the workshop and utilize the case studies to illustrate key points regarding the importance of decision context and confidence in the methods for their successful application in making key public health decisions. This information is organized according to decision context in Table 1, which is meant to be illustrative in that it presents a range of contexts, issues, and examples but does not represent an exhaustive list.

Decision Context

A key theme emerging from the workshop was that decision context—that is, the nature of the problem, the reason for addressing the problem, and the policy or public health impact of the decision—is crucial to the acceptance of new types of data. Literature in the history, philosophy, and sociology of science emphasizes that in order to make effective and socially responsible use of new scientific methods and tools such as those currently under consideration in the fields of toxicology, exposure science, and risk assessment, it is crucial to consider the decision context in which they will be employed, including the

Table 1. Use of high-throughput screening (HTS) technologies across different decision contexts.

Decision context	Problem	Possible approach	Barriers	Examples	Comment
Hazard identification	*In vivo* testing is resource and time intensive, many data gaps	HTS tests rapidly assess biological activities relating to *in vivo* toxicity	How well anchored to *in vivo*? Rate of false negatives?	Mutagenicity test batteries; Estrogen receptor model (Case Study 1)	High confidence HTS approaches may replace whole animal testing, whereas lower confidence HTS may augment chemical screening; confidence aided by few false negatives and HTS linkage to AOP
Chemical prioritization	Large chemical testing backlog	Exposure predictions, biomonitoring, or HTS used to identify chemicals of highest concern	Not all end points have HTS equivalent; IVIVE has uncertainties	HTS + IVIVE + ExpoCast evaluation of 3925 chemicals (Sipes et al. 2017)	Rapid bioactivity and exposure screens have limitations and extrapolation to *in vivo* dose is imprecise; useful for prioritization but not exclusion of chemicals from further evaluation
Environmental media quality	Targeted analysis of chemicals may miss chemicals and interactive effects	HTS bioassays another way to assess environmental quality	Environmental standards not established for bioassay-based approaches	Water quality assessment (Case Study 2)	HTS can show need for and effectiveness of control technologies even if HTS results not clearly linked to *in vivo* end points; need to develop consistent HTS protocols for media screening
Consumer products ingredient	High concern chemical in products needs replacement	HTS testing can help fill data gaps and support alternatives assessments	HTS may miss end points and human vulnerabilities; may replace one adverse property with others	California Safer Consumer Products Program (Case Study 3)	HTS has an increasing role along with SAR, exposure forecasting and *in vivo* results in alternative assessment by increasing data available on alternatives
Emergency response	Time-critical decisions involving chemicals during emergencies	HTS exposure and toxicity approaches can screen agents for health and environmental harm	HTS limited with respect to predicting chemical fate and effects in environment	Dispersants in *Deepwater Horizon*; MCHM spill (Case Study 4)	HTS has a key role where data gaps impair rapid decision-making, but limitations can result in potentially missed hazards; testing should ideally be done prior to entry in environment

Note: AOP, adverse outcome pathway; HTS, high-throughput screening; IVIVE, *in vitro*–to–*in vivo* extrapolation; MCHM, 4-methylcyclohexanemethanol; SAR, structure–activity relationship.

ultimate purposes for which they are going to be used (Douglas 2009; Elliott and McKaughan 2014; Fernández Pinto and Hicks 2019). For example, in situations where the costs of false-positive (Type I) and/or false-negative (Type II) errors are relatively low, it may be justifiable to rapidly employ novel methods even if there is uncertainty about their reliability (Douglas 2009). In contrast, when both decision stakes and uncertainties are high, more public engagement may be needed before introducing new methods or tools (Funtowicz and Ravetz 1993).

Another aspect to consider, in addition to the cost of potentially inaccurate information, is the cost of having minimal or no information. In such a case, even newly developed but not well-vetted methods may aid decision-making. In our view, these cases may be particularly challenging but also point to the greatest needs and opportunities for using new methods and fostering their further evolution. Addressing these cases may require that multiple interested and affected parties clarify the types of information that would be of greatest use while considering the decision timeframe and anticipated human exposures. Various disciplinary perspectives are needed to answer those questions along with considering the degree of confidence needed for a particular decision (Elliott 2017; Intemann 2015; Lacey 1999). Another key challenge is data integration and interpretation for high-throughput chemical assessment findings because the statistical and machine learning methods for analyzing and reporting this information need optimization for use by researchers and decision-makers (Kosnik et al. 2019).

The needs in this area are highlighted by the recent amendments to the Toxic Substances Control Act (TSCA) (U.S. EPA 2016, 2018). These amendments encourage the use of data from alternative test methods and strategies. A key aspect of the amended TSCA is developing strategic approaches to integrate emerging data into decision-making, an activity that will provide examples we believe can help the risk assessment community at large understand the utility of these methods for not only chemicals management policy but for a range of applications. In summary, we find that when designing and implementing new methods for determining hazard, exposure, and risk for chemicals in the environment, the decision context, and the values inherent in that context (e.g., improved protection of health and the environment), must be considered early in the process to ensure maximum utility for decision-making.

Values in Decision-Making

The manner in which the new data are applied will be embedded with values associated with the level of precaution and health protection desired or required. Making these value judgments transparent and consistent with existing policy will increase confidence that their introduction will not shift decision-making toward less protection. For example, when prioritizing chemicals for further study for a particular biological outcome (Case Study 1 below), positive results (i.e., results that indicate potential harm) in relevant bioassays could be used to identify chemicals of concern, whereas negative results (i.e., results that indicate a lack of potential harm) are not sufficient to conclude a lack of concern given the limitations of current *in vitro* methods to simulate *in vivo* metabolism or predict effects in different tissues and across different life stages. If the goal of the decision is to evaluate the quality of environmental media (Case Study 2 below), screening for bioactivity in a wide range of assays may be useful to show, in a nontargeted fashion, contamination that would be of value to mitigate even though the public health implications may not be clear. If the goal of the decision is to identify greener/safer alternatives (Case Study 3), a range of forecasting tools (structural information, exposure

modeling, high-throughput toxicology testing) could be used along with traditional toxicology results to ensure that one chemical of concern is not replaced with another harmful chemical. If the goal of the decision is to inform emergency response for novel chemicals released into the environment (Case Study 4), rapidly deployed computational and *in vitro* screening, and short-term targeted testing in vertebrate and invertebrate systems could inform hazard-based decision-making.

In each of these decision contexts, we find that a key underlying value is to increase the information available and the use of this information in a public health protective manner even if the new information cannot at this point be converted into a quantitative prediction of population health risk. The current system of environmental chemical evaluations has been criticized as favoring the protection of chemicals over people (Krimsky 2017) due, in part, to the slow pace at which existing chemicals have been evaluated and the dearth of information collected about new industrial chemicals that enter into the market. This landscape can potentially be reshaped by the large amounts of data that can be generated on chemicals by helping to fill data gaps and eliminate the inherent default of no data equals no risk (NRC 2009).

We consider it essential that decision-makers apply new tools in ways that serve a protective role. Toward this end, tools should be designed to be robust screens that minimize false negatives (Type II errors) while permitting higher levels of false positives (Type I errors) in order to buffer against harms caused by toxic chemicals in the environment, potentially for long periods of time (Cranor 2017). The application of emerging tools that do not yet cover all relevant aspects of biology could result in the undesirable outcome of concluding that a chemical poses a low or no hazard when, in fact, it poses a potentially significant hazard that is not captured by the assays (false negative). For example, recent evaluations have noted that the existing platforms in ToxCast™ do not cover all the key characteristics of carcinogens, potentially missing chemicals that may have the potential to cause cancer through other pathways (Iyer et al. 2019; Guyton et al. 2018). Thus, the potential for false negatives must be carefully considered because such outcomes could have far-ranging and long-lasting adverse impacts.

The use of new technologies in decision-making must also be protected from perpetuating or amplifying existing health disparities. Extensive literature documents the disproportionate health and environmental hazards faced by many communities of color, low-income communities, and Indigenous communities (Bullard 2000; Burwell-Naney et al. 2013; Lee and Mohai 2012; American Lung Association 2001; Davis et al. 2016; Seabury et al. 2017; Zota and Shamasunder 2017). Disproportionate exposures to toxic chemicals coupled with exposures to social hazards (e.g., poverty) can result in an amplification of harm faced by these communities (Gee and Payne-Sturges 2004). When emerging tools provide false-negative results, or slow down the process by which chemicals are evaluated (i.e., simply serving as an additional step rather than making chemical evaluation more efficient), the added burden of exposure and hazard may be disproportionately experienced by Environmental Justice communities.

In summary, we point out that the manner in which toxicology data, in general, and emerging new data streams, in particular, are evaluated and applied have embedded values regarding the level of public health protection. These values should be made transparent and consistent with the goals of improving public health and environmental protection through more rapid and efficient chemical screening, decreasing data gaps, and ensuring that the potential for false-negative outcomes is minimized to the extent possible.

Instilling Confidence in New Methods

There are many steps in the development of novel methods and in instilling confidence in their applicability and reliability. These steps begin with recognizing a need for change, either because the existing methods are deficient or new technologies present opportunities, such as efficiency, reduced cost, or improved decision-making ability. The next step involves a scoping exercise, where the applications and decision contexts are considered and the research effort is focused on a fit-for-purpose level of evaluation (i.e., the data generated by the method fits the needs of the particular decision context and value structure under which it will be used). Scoping should be followed by technical aspects of method development and validation/evaluation through which predictivity, reliability, and responsiveness to the range of human variability can be assessed. The utility and applicability of these methods must then be considered relative to the goals and values of method development. Widespread adoption will require both matching the available data to the data needed for decision-making and making sure that the large amounts of data that can be rapidly generated across a range of different bioassay systems can be distilled and interpreted for the needed context. This step, data integration and interpretation, should follow agreed-upon protocols so that the process is transparent to decision-makers as well as to the general public. Developing clear links between *in vitro* assays and *in vivo* end points in terms of showing the reliability and predictive capability of the new methods will increase confidence but may be a data integration and interpretation challenge given that many *in vitro* assays target elements of networks that support multiple biological outcomes. However, as illustrated in the case studies, the need for *in vitro* assays to be linked to particular health outcomes varies across decision contexts and so the level of data synthesis and interpretation that is needed will also vary.

In summary, increased transparency in the new methods, data collection, and analysis through agreed-upon protocols and the development of peer review publications, as well as ample opportunities for meaningful multi-directional risk communication, are critical aspects of confidence building.

Case Examples

Four case studies were considered at the NASEM Paradigm Shift meeting (NASEM 2018) to highlight the information and confidence needs for emerging data across a range of decision contexts. These case studies are used in this commentary to illustrate the useful application of emerging methods and to further explore how their design and development could be aligned with the needs of specific decision contexts and with the overall goal of improved health protection.

1. In Vitro *Battery for Screening Estrogenicity of Potentially Endocrine Active Chemicals*

The U.S. EPA Endocrine Disruptor Screening Program (EDSP) was established in response to section 408(p) of the Federal Food, Drug, and Cosmetics Act [21 U.S.C. 346a(p)(1)] and section 1457 of the Safe Drinking Water Act (42 U.S.C. 300j-17) to develop a screening program to identify chemicals that could interfere with the endocrine system. Based upon recommendations of the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) and the U.S. EPA Administrator, the agency created a two-tier process for screening the approximately 10,000 chemicals that have widespread exposure in the environment with potential estrogenic, androgenic, and/or thyroid disrupting activity.

Beginning in 2011, the U.S. EPA EDSP began a multiyear process to integrate the use of new methods for the prioritization and screening of chemicals with potential endocrine activity (U.S. EPA 2015a). This process involved a pivot away from a framework that relied on the use of low-to-medium–throughput assays toward a framework that incorporated batteries of high-throughput *in vitro* assays to screen chemicals for endocrine activity.

As one of the first steps in the pivot, the agency compiled a battery of *in vitro* assays and developed a mechanistic interpretation of the battery's results to replace three existing tests—an *in vitro* estrogen receptor (ER) binding assay, an *in vitro* ER transactivation assay, and an *in vivo* uterotrophic assay—for screening estrogenic potential of EDSP-relevant chemicals (Browne et al. 2015). This pathway-based approach required an extensive literature review to identify reference chemicals, testing the predictive validity of approaches, and gathering peer review and public comment. The effort required strategic collaboration and cooperation with and among the regulated community, regulators, and stakeholders, and ultimately resulted in the acceptance of the ER battery as a replacement under EDSP. Efforts are currently underway for the adoption of similar approaches for androgen activity, thyroid pathways, and steroidogenesis (Erickson 2018).

For this particular decision context (i.e., the hazard determination for a specific end point), the adoption of new tools benefitted from an existing gold standard of ER screening assays and a learning set of chemicals tested in those assays whose results could be compared with results from high-throughput assays. Further advantage was taken of the relatively well-understood pathway for genomic ER-mediated gene transcription and its resulting effects.

We find this case example points out two factors that can help build confidence in emerging methods: *a*) the existence of an extensive comparison data set of *in vivo* results, and *b*) the ability to construct a battery of high-throughput assays patterned upon a mechanism-based adverse outcome pathway (AOP). These factors created a high level of confidence in the ability of the new methods to screen for estrogenic activity with increased efficiency.

The benefit to public health conferred by the application of these tools remains undetermined until there is a body of decisions based upon the new high-throughput estrogen screen. The potential exists for these tools to aid in the rapid identification and regulation of endocrine disrupting chemicals, and the methods have the capability to generate estimates of sensitivity and specificity that could build confidence in the tools. However, the extent by which their use may enhance health protection will require future evaluation.

2. In Vitro *Screens to Assess Water Quality*

Chemicals in sewage effluent, surface water, and drinking water can be extremely complex mixtures of regulated and unregulated substances. Increasing public awareness of water-related issues, particularly for emerging substances such as per- and polyfluoroalkyl substances (PFASs), heighten the need for assessing and cleaning water sources. Complex chemical mixtures such as those that can be found in water highlight the utility of expanding beyond one-chemical-at-a-time evaluations to those that incorporate larger universes of chemical constituents (Cizmas et al. 2015). The ability to assess mixtures across a wide range of relevant end points, as well as the ability to evaluate the effectiveness of different treatment strategies for water purification, make water quality assessment a potentially rich application of emerging technologies for the protection of public health.

For a decision context that involves determining the relative quality of various water sources or the quality of remediation

technologies, testing has shown water quality improvements when comparing wastewater treatment effluent to the receiving river water and to treated drinking water. Bioassay-based approaches have also been used to assess the residual toxicity associated with drinking water contaminant destruction technologies such as advanced oxidation (Escher et al. 2013). This approach is most useful when there is some water quality reference point (e.g., a pristine source of water) against which other samples can be judged. These tools may also help in building consumer confidence in the industrial and domestic uses of reclaimed/recycled water (Leusch et al. 2014).

This case example illustrates that even when there is limited understanding of the *in vivo* relevance of high-throughput testing, such testing can be a useful aid to decisions regarding best available control technologies and the acceptability of media quality relative to some agreed-upon reference material (e.g., a pristine water source). However, the presence of biological activity in a water sample as an indicator of water quality impairment may be difficult to apply because there are no standards or regulatory criteria for such impairment to date. Even with these limitations, we believe that these emerging tools can assist in remediation-related decision-making (e.g., filtration technologies utilized by water utilities, analysis of effluent from releasing facilities).

3. Evaluation of Alternatives for Consumer Products

The State of California is engaged in an effort to drive the development of safer consumer products (DTSC 2013). The California Safer Consumer Products Program selects combinations of products and chemicals that have the potential for exposure and hazard and requires manufacturers to systematically evaluate functional alternatives. The DTSC Candidate Chemical list contains approximately 2,500 chemicals and several large chemical classes (CalSAFER 2019). Chemicals under evaluation in certain product categories include several organohalogen flame retardants, methylene diisocyanate, methylene chloride, PFASs, *N*-methyl-2-pyrrolidone, and alkylphenol polyethoxylates (Solomon et al. 2018).

A key decision context is whether safer alternatives can be found for chemicals of concern. Although no alternatives analyses have yet been completed, DTSC expects to receive extensive evidence from new data streams for emerging chemicals under consideration as potential functional alternatives (Smith et al. 2019; Malloy et al. 2017). For example, prediction of physicochemical properties, quantitative structure–activity relationship (QSAR) models, *in vitro* data, and data from nonmammalian systems (e.g., Zebrafish) could prove useful in screening alternatives to identify those that are predicted to have the least hazard and/or exposure potential.

We find this case example illustrates the principle that decision contexts that require rapid screening of many chemical alternatives with a highly variable amount of toxicology information are an especially promising area for near-term use of newer methods. In such a context, any evidence of potential hazard could be included in a decision matrix that includes qualitative factors such as level of confidence in a particular effect and comparability of testing across candidate chemicals. Although these approaches hold the promise of protecting public health and the environment by identifying safer alternatives, we caution that lack of evidence of hazard does not necessarily mean the preferred alternative is completely safe but, rather, that it does not show evidence of known hazard potential.

4. Environmental and Health Emergencies

Understanding the health and environmental risks posed by data-poor chemicals released into the environment is an important decision context into which emerging data streams may play a significant role. Such methods were used by the U.S. EPA to rapidly screen multiple dispersant chemicals for potential aquatic toxicity and endocrine disruption potential during the 2010 *Deepwater Horizon* oil spill in the Gulf of Mexico (Judson et al. 2010). They were also used in the rapid response to the 2014 chemical spill into the Elk River in West Virginia (NTP 2019). In both situations, chemicals with limited toxicity information needed to be evaluated rapidly (in days or weeks) and emerging methods were employed to evaluate the potential impacts to humans, animals, and ecosystems.

In the *Deepwater Horizon* example, *in vitro* high-throughput assays were used to evaluate eight commercial dispersants that were considered for use in the massive oil spill cleanup activities (Judson et al. 2010). Evaluation focused on estrogenic and androgenic activity, other biological pathways, and mammalian cell cytotoxicity. The information generated was used to help select a specific dispersant that was used on the spill, Corexit® 9500. In the years since the *Deepwater Horizon* disaster, studies have identified potential human health and ecosystem impacts associated with the use of dispersants (Alexander et al. 2018; DeLeo et al. 2016; McGowan et al. 2017; Ramesh et al. 2018; Zhang et al. 2013), illustrating the complexities and trade-offs decision-makers face when dealing with unknowns in emergency situations, and where reliance on some rather than no toxicity information may still fall short in the protection of public health and the environment.

In the example of the 2014 contamination of the Elk River with a mixture of chemicals used to clean coal, primarily 4-methylcyclohexanemethanol (MCHM), health authorities urgently required information in order to advise the nearly 300,000 residents of the Charleston, West Virginia, area about the safety of their drinking water. At the time of the spill, West Virginia American Water issued a do-not-drink order due to a lack of information about the chemical. The Centers for Disease Control and Prevention (CDC) initially identified a screening level value of 1 ppm based on information available at the time.

In the months following the spill, the CDC nominated MCHM and the other chemicals involved in the Elk River contamination event for further investigation by the National Toxicology Program (NTP 2014). Over the following year, the NTP performed a range of studies that evaluated the potential impacts of MCHM and other chemicals on development and growth, skin irritation and hypersensitivity, behavior, DNA mutation and genetic damage, and molecular effects on biological processes. Studies involved *in vivo* and *in vitro* experiments, including transcriptomic studies in rats and developmental studies in *Caenorhabditis elegans* and Zebrafish. Based upon their test results, the NTP concluded that the data supported the adequacy of the drinking-water screening level recommended by the CDC at the time of the spill (NTP 2016).

In this example, the follow-up tests performed by the NTP demonstrate that new tools could help support decisions about chemicals in the environment. Although the NTP studies were helpful in retrospectively evaluating a decision, emergency situations may require much more rapid evaluation of potential health and environmental impacts. In such circumstances, we emphasize that it is critically important to interpret information from rapid assessments in ways that place high value on health protection and err on the side of precaution. A far better solution would be to conduct testing prior to commercial use or storage of a chemical to ensure that the information is complete and immediately available in the event of an environmental release. This case example highlights the potential to develop health and environmental fate/exposure data to make rapid predictions of chemical effects upon release into the environment during an emergency.

Conclusions

The ability to readily assess the hazard, exposure, and risk associated with chemicals is a pressing need for decision-makers across a wide range of decision contexts. The utility of the information generated by new tools for the range of decision contexts and levels of health protection being targeted is a critical consideration in the evolution of these methods. Although the beginnings of a paradigm shift may be slow and incremental, increased understanding of the ways in which the application of new tools could increase—or potentially decrease—human health and environmental protections can assist decision-makers in calibrating the speed of incorporating new methodologies and data streams.

We find that the variety of assays that can be conducted and the large amounts of data becoming available necessitates a strategic approach that focuses on the needs of a particular scenario or decision context when designing, synthesizing, and communicating the results of these emerging techniques. The values embedded in their application need to be transparent and consistent with the underlying goals of public health and environmental protection.

As illustrated by the four case studies, the manner in which the data can be of use to decision-makers can vary widely, and new applications will likely present themselves over time. Although these cases demonstrate how new methods can assist decision-making, we point out that they also provide cautionary tales and illustrate the need for researchers, various stakeholders, and decision-makers to agree on the data needed to build confidence in using new methods for specific purposes. Confidence-building measures can include the measurement of performance against gold standard assays, especially with respect to limiting the number of false negatives; the availability of prototype chemicals for read-across approaches; and the linkage of high-throughput assays to AOPs. As these evidence streams develop further, they have the potential to play an increasing role in screening for hazardous properties, prioritizing chemicals for further testing, identifying safer alternatives, assessing environmental media, improving emergency response, and, overall, providing greater protection of public health and the environment.

Acknowledgments

References

Alexander M, Engel LS, Olaiya N, Wang L, Barrett J, Weems L, et al. 2018. The Deepwater Horizon oil spill coast guard cohort study: a cross-sectional study of acute respiratory health symptoms. Environ Res 162:196–202, PMID: 29331799, https://doi.org/10.1016/j.envres.2017.11.044.

American Lung Association. 2001. Urban air pollution and health inequities: a workshop report. Environ Health Perspect 109(suppl 3):357–374, PMID: 11427385, https://doi.org/10.2307/3434783.

Browne P, Judson RS, Casey WM, Kleinstreuer NC, Thomas RS. 2015. Screening chemicals for estrogen receptor bioactivity using a computational model. Environ Sci Technol 49(14):8804–8814, PMID: 26066997, https://doi.org/10.1021/acs.est.5b02641.

Bullard RD. 2000. *Dumping in Dixie: Race, Class, and Environmental Quality.* Boulder, CO: Westview Press.

Burwell-Naney K, Zhang H, Samantapudi A, Jiang C, Dalemarre L, Rice L, et al. 2013. Spatial disparity in the distribution of Superfund sites in South Carolina: an ecological study. Environ Health 12:96, PMID: 24195573, https://doi.org/10.1186/1476-069X-12-96.

CalSAFER. 2019. Search DTSC Candidate Chemical Database. https://calsafer.dtsc.ca.gov/Chemical/search.aspx [accessed 26 June 2019].

Cizmas L, Sharma VK, Gray CM, McDonald TJ. 2015. Pharmaceuticals and personal care products in waters: occurrence, toxicity, and risk. Environ Chem Lett 13(4):381–394, PMID: 28592954, https://doi.org/10.1007/s10311-015-0524-4.

Cranor CF. 2017. *Tragic Failures: How and Why We are Harmed by Toxic Chemicals.* New York, NY: Oxford University Press.

Davis HT, Aelion CM, Liu J, Burch JB, Cai B, Lawson AB, et al. 2016. Potential sources and racial disparities in the residential distribution of soil arsenic and lead among pregnant women. Sci Total Environ 551–552:622–630, PMID: 26897405, https://doi.org/10.1016/j.scitotenv.2016.02.018.

DeLeo DM, Ruiz-Ramos DV, Baums IB, Cordes EE. 2016. Response of deep-water corals to oil and chemical dispersant exposure. Deep-Sea Res Part 2 Top Stud Oceanogr 129:137–147, https://doi.org/10.1016/j.dsr2.2015.02.028.

Douglas HE. 2009. *Science, Policy, and the Value-Free Ideal.* Pittsburgh, PA: University of Pittsburgh Press.

DTSC (Department of Toxic Substances Control). 2013. 22 CCR § 69506.2-.8. 2013. Safer Consumer Product Regulations. R-2011-02. https://dtsc.ca.gov/regs/safer-consumer-product-regulations/ [accessed 27 June 2019].

Elliott K. 2017. *A Tapestry of Values: An Introduction to Values in Science.* New York, NY: Oxford University Press.

Elliott KC, McKaughan DJ. 2014. Nonepistemic values and the multiple goals of science. Philos Sci 81(1):1–21, https://doi.org/10.1086/674345.

Erickson BE. 2018. Endocrine receptor assays go fast track. Chem Eng News 96(11):22–82.

Escher BI, van Daele C, Dutt M, Tang JYM, Altenburger R. 2013. Most oxidative stress response in water samples comes from unknown chemicals: the need for effect-based water quality trigger values. Environ Sci Technol 47(13):7002–7011, PMID: 23432033, https://doi.org/10.1021/es304793h.

Fernández Pinto M, Hicks DJ. 2019. Legitimizing values in regulatory science. Environ Health Perspect 127(3):35001, PMID: 30870036, https://doi.org/10.1289/EHP3317.

Funtowicz SO, Ravetz JR. 1993. Science for the post-normal age. Futures 25(7):739–755, https://doi.org/10.1016/0016-3287(93)90022-L.

Gee GC, Payne-Sturges DC. 2004. Environmental health disparities: a framework integrating psychosocial and environmental concepts. Environ Health Perspect 112(17):1645–1653, PMID: 15579407, https://doi.org/10.1289/ehp.7074.

Guyton KZ, Rusyn I, Chiu WA, Corpet DE, van den Berg M, Ross MK, et al. 2018. Application of the key characteristics of carcinogens in cancer hazard identification. Carcinogenesis 39(4):614–622, PMID: 29562322, https://doi.org/10.1093/carcin/bgy031.

Intemann K. 2015. Distinguishing between legitimate and illegitimate values in climate modeling. Eur J Philos Sci 5(2):217–232, https://doi.org/10.1007/s13194-014-0105-6.

Iyer N, Pham N, Marty M, Sandy M, Solomon G, Zeise L. 2019. An integrated approach using publicly available resources for identifying and characterizing chemicals of potential toxicity concern: proof-of-concept exercise with chemicals that affect cancer pathways. Toxicol Sci 169(1):14–24, PMID: 30649495, https://doi.org/10.1093/toxsci/kfz017.

Judson RS, Martin MT, Reif DM, Houck KA, Knudsen TB, Rotroff DM, et al. 2010. Analysis of eight oil spill dispersants using rapid, in vitro tests for endocrine and other biological activity. Environ Sci Technol 44(15):5979–5985, PMID: 20602530, https://doi.org/10.1021/es102150z.

Kosnik MB, Planchart A, Marvel SW, Reif DM, Mattingly CJ. 2019. Integration of curated and high-throughput screening data to elucidate environmental influences on disease pathways. Comput Toxicol 12:100094, PMID: 31453412, https://doi.org/10.1016/j.comtox.2019.100094.

Krimsky S. 2017. The unsteady state and inertia of chemical regulation under the US Toxic Substances Control Act. PLoS Biol 15(12):e2002404, PMID: 29252997, https://doi.org/10.1371/journal.pbio.2002404.

Lacey H. 1999. *Is Science Value Free? Values and Scientific Understanding.* London, UK: Routledge.

Lee S, Mohai P. 2012. Environmental justice implications of brownfield redevelopment in the United States. Soc Nat Resour 25(6):602–609, https://doi.org/10.1080/08941920.2011.566600.

Leusch FDL, Khan SJ, Gagnon MM, Quayle P, Trinh T, Coleman H, et al. 2014. Assessment of wastewater and recycled water quality: a comparison of lines of evidence from in vitro, in vivo and chemical analyses. Water Res 50:420–431, PMID: 24210511, https://doi.org/10.1016/j.watres.2013.10.056.

Malloy T, Zaunbrecher V, Beryt E, Judson R, Tice R, Allard P, et al. 2017. Advancing alternatives analysis: the role of predictive toxicology in selecting safer chemical products and processes. Integr Environ Assess Manag 13(5):915–925, PMID: 28247928, https://doi.org/10.1002/ieam.1923.

McGowan CJ, Kwok RK, Engel LS, Stenzel MR, Stewart PA, Sandler DP. 2017. Respiratory, dermal, and eye irritation symptoms associated with Corexit™ EC9527A/EC9500A following the *Deepwater Horizon* Oil Spill: findings from the GuLF STUDY. Environ Health Perspect 125(9):097015, PMID: 28934097, https://doi.org/10.1289/EHP1677.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2018. *Understanding Pathways to a Paradigm Shift in Toxicity Testing and Decision-Making: Proceedings of a Workshop—in Brief.* Washington, DC: National Academies Press.

NRC (National Research Council). 2009. *Science and Decisions: Advancing Risk Assessment*. Washington, DC: The National Academies Press.

NTP (National Toxicology Program). 2014. Nomination Summary for West Virginia Elk River chemical spill (N21408). https://ntp.niehs.nih.gov/testing/noms/search/summary/nm-n21408.html?utm_source=direct&utm_medium=prod&utm_campaign=ntpgolinks&utm_term=nm-n21408 [accessed 30 November 2019].

NTP. 2016. NTP Research Program on Chemicals Spilled into the Elk River in West Virginia: Final Update. https://ntp.niehs.nih.gov/ntp/research/areas/wvspill/wv_finalupdate_july2016_508.pdf [accessed 26 June 2019].

NTP. 2019. West Virginia Chemical Spill: NTP Research Response and Findings. https://www.niehs.nih.gov/health/materials/west_virginia_chemical_spill_508.pdf [accessed 26 June 2019].

Ramesh S, Bhattacharya D, Majrashi M, Morgan M, Prabhakar Clement T, Dhanasekaran M. 2018. Evaluation of behavioral parameters, hematological markers, liver and kidney functions in rodents exposed to Deepwater Horizon crude oil and Corexit. Life Sci 199:34–40, PMID: 29474811, https://doi.org/10.1016/j.lfs.2018.02.028.

Seabury SA, Terp S, Boden LI. 2017. Racial and ethnic differences in the frequency of workplace injuries and prevalence of work-related disability. Health Aff (Millwood) 36(2):266–273, PMID: 28167715, https://doi.org/10.1377/hlthaff.2016.1185.

Sipes NS, Wambaugh JF, Pearce R, Auerbach SS, Wetmore BA, Hsieh JH, et al. 2017. An intuitive approach for predicting potential human health risk with the Tox21 10k library. Environ Sci Technol 51(18):10786–10796, PMID: 28809115, https://doi.org/10.1021/acs.est.7b00650.

Smith MN, Cohen Hubal EA, Faustman EM. 2019. A case study on the utility of predictive toxicology tools in alternatives assessments for hazardous chemicals in children's consumer products. J Expo Sci Environ Epidemiol, PMID: 31501452, https://doi.org/10.1038/s41370-019-0165-y.

Solomon G, Reynolds P, Hoang A. 2018. *California's Green Chemistry Initiative at Age 10: An Evaluation of its Progress and Promise*. Oakland, CA: Public Health Institute. http://www.phi.org/resources/?resource=california-green-chemistry-report [accessed 26 June 2019].

U.S. EPA (U.S. Environmental Protection Agency). 2015a. Endocrine Disruptor Screening Program for the 21st Century (EDSP21 Work Plan). https://www.epa.gov/sites/production/files/2015-07/documents/edsp21_work_plan_summary_overview_final.pdf [accessed 26 June 2019].

U.S. EPA. 2015b. National ambient air quality standards for ozone. Fed Reg 80(206):65292–65468. RIN 2060-AP38. https://www.govinfo.gov/content/pkg/FR-2015-10-26/pdf/2015-26594.pdf [accessed 26 June 2019].

U.S. EPA. 2016. The Frank R. Lautenberg Chemical Safety for the 21st Century Act. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/frank-r-lautenberg-chemical-safety-21st-century-act [accessed 26 June 2019].

U.S. EPA. 2018. Alternative Test Methods and Strategies to Reduce Vertebrate Animal Testing. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/alternative-test-methods-and-strategies-reduce [accessed 14 November 2018].

Zhang Y, Chen D, Ennis AC, Polli JR, Xiao P, Zhang B, et al. 2013. Chemical dispersant potentiates crude oil impacts on growth, reproduction, and gene expression in *Caenorhabditis elegans*. Arch Toxicol 87(2):371–382, PMID: 22990136, https://doi.org/10.1007/s00204-012-0936-x.

Zota AR, Shamasunder B. 2017. The environmental injustice of beauty: framing chemical exposures from beauty products as a health disparities concern. Am J Obstet Gynecol 217(4):418e1–418e6, PMID: 28822238, https://doi.org/10.1016/j.ajog.2017.07.020.